**Examine the Impact of Nonnormality on Parameter Estimates in the Multidimensional (Bi-factor) graded response model Manuscript**

**Model and Definition**

**Model Description**

Item 1

…

Item 10

Item 11

…

Item 20

Item 21

…

Item 30

Bifactor-IRT

In this study, we focus on the association between the constructs of psychometric tests, in which items are always linked to multiple latent traits. Hence, we extend the unidimensional GRM model to the multidimensional GRM (MGRM). in GRM is changed to , a vector of length H representing the latent traits of interest, the multidimensional version of probability equation is

, (4)

In which . This parameterization with D=1 is consistent with flexMIRT’s (Cai, 2015) default parameterization, serving as the “intercept”.

**Estimation of Non-Normality**

Pearson Type 2 skewness is calculated by:

(8)

Based on pre-text result, we choose G1 as the skewness estimation method. Instead of comparing mean of skewness of predicted theta with setting skewness (0, 0.4, 0.8), we compute the mean of skewness of generated theta (real theta) to calculate the bias between skewness of predicted theta and skewness of generated theta.

**Simulation**

**Design**

Table 1 Simulation Design

|  |  |  |
| --- | --- | --- |
| Manipulated parameters | Number of levels | Values of levels |
| **Data Structure** |  |  |
| Sample size (N) | 3 | N = 250, 500, 1000 |
| Item per Factor (J) | 2 | J = 5, 10 |
| Factor (F) | 3 | F = 2, 3, 4 |
|  |  |  |
| **Skewness of Theta** |  |  |
| Skewness on general factor (GF) | 3 | GF= 0, 0.4, 0.8 |
| Skewness on specific/group factor (SF) | 3 | SF = 0, 0.4, 0.8 |

**Item parameter**

ag, was set as [1.1, 2.8].

aS was set as [0, 1.5]

This study randomly generated b1[−2, −0.67], b2[−0.67, 0.67], b3[0.67, 2] for three thresholds to distinguish the possibilities of choosing each item.

**Person parameter**

There are three levels of skewness (normal=0, moderate=0.4, severe=0.8) were manipulated for each general factor and specific factors. Although general latent trait and specific latent traits have different discrimination ag, no previous research focused on different algorithms’ performances on recovery of θg and θs. Here, different severity of skewness will be applied on two latent traits of single individual. There were nine combinations of non-skewed, moderately skewed, and severely skewed.

|  |  |  |
| --- | --- | --- |
| Normal (GF) – Normal (SF) | Normal (GF) – Moderate (SF) | Normal (GF) – Severe (SF) |
| Moderate (GF) – Normal (SF) | Moderate (GF) – Moderate (SF) | Moderate (GF) – Severe (SF) |
| Severe (GF) – Normal(SF) | Severe (GF) – Moderate (SF) | Severe (GF) – Severe (SF) |

In package of R, “sn”, used to simulate theta, latent trait of individuals, the skewness is ranged from 0 to 0.999. This study set three level of skewness as 0, 0.4, 0.8, separately. Because we only focus the difference of estimation of θg on general factor and specific factor in various conditions, all latent traits on specific factors (θsk) are set equally.

**Estimation**

A fully crossed design for all these manipulated factors yielded a total of 162 conditions, each of which was replicated 30 times using packages and code written in R (R Core Team, 2021). Full-information maximum likelihood estimation in FlexMIRT is the only method utilized in this paper, the first step of the whole research. To increase the convergence speed in FlexMIRT, the number of integration quadrature points was reduced to 21 (default is 49), and the range was set to -3.5 to 3.5, which is also applied in Wang’s research (2018). Because FlexMIRT only offer cjk instead of bjk, we use bjk = -cjk/aj to transit cjk to bjk.

***Replications***

In a pilot study conducted in which the simulation procedure was implemented on 36 conditions for 50 replications, Wang (2018) found that the recovery indices stabilized after 30 replications. According to this result, we decided to replicate each of the 162 conditions in the study for 30 times. Means and standard deviations of each dependent variable were computed across replications.

***Evaluation criteria***

The relative bias is estimated for all the parameters of model (ag, as, b1, b2, b3) as,

(9)

In which, is the predicted parameters () across valid replications and is the real parameters (agj, asj, b1j, b2j, b3j) simulated at the beginning of this research. In the , *j* is the order of items in each condition, from 1 to N, N calculated by number of items in each specific factor multi by number of specific factors. In equation, the number of replications for each condition is from 1 to M. In this study, M is fixed at 30.

The RMSE of

. (10)

The RMSE depends on the balance between the bias and standard deviation or variance. To estimate how skewness of density of real persons’ latent traits θg and θs straightforwardly, this study will estimate the variance of these parameters, instead of RMSE.

A five-way analysis of variance (ANOVA) for Bias, RMSE, and Variance were employed to compare and summarize the results of this simulation study.